An Integrated Genetic Map of the African Human Malaria Vector Mosquito, *Anopheles gambiae*

Liangbiao Zheng,*,† Mark Q. Benedict,‡ Anton J. Cornel,‡ Frank H. Collins‡ and Fotis C. Kafatos*,†

*Department of Molecular and Cellular Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusets 02138, †European Molecular Biology Laboratory, 69117 Heidelberg, Germany and ‡Malaria Branch F(12), Centers for Disease Control, Atlanta, Georgia 30333

ABSTRACT

We present a genetic map based on microsatellite polymorphisms for the African human malaria vector, *Anopheles gambiae*. Polymorphisms in laboratory strains were detected for 89% of the tested microsatellite markers. Genotyping was performed for individual mosquitoes from 13 backcross families that included 679 progeny. Three linkage groups were identified, corresponding to the three chromosomes. We added 22 new markers to the existing *X* chromosome map, for a total of 46 microsatellite markers spanning a distance of 48.9 cM. The second chromosome has 57 and the third 28 microsatellite markers spanning a distance of 72.4 and 93.7 cM, respectively. The overall average distance between markers is 1.6 cM (or 1.1, 1.2, and 3.2 cM for the *X*, second, and third chromosomes, respectively). In addition to the 131 microsatellite markers, the current map also includes a biochemical selectable marker, *Dieldrin resistance* (*Dl*), on the second chromosome and five visible markers, *pink-eye* (*p*) and *white* (*w*) on the *X*, *collarless* (*c*) and *lunate* (*lu*) on the second, and *red-eye* (*r*) on the third. The cytogenetic locations on the nurse cell polytene chromosomes have been determined for 47 markers, making this map an integrated tool for cytogenetic, genetic, and molecular analysis.

*A*NOPHELES *gambiae* is the principal vector for human malaria in Africa, causing an estimated 200 million clinical cases and more than one million deaths annually. Considerable interest is now focused on a possible malaria control strategy based on replacement of field mosquito populations by strains refractory to parasites. Although a number of obstacles currently exist for such a strategy, it is encouraging that refractoriness is not an uncommon mosquito trait and that it appears to involve a small number of genetic loci (SEVERSON 1994). Thus, a few selection steps in a laboratory population of *A. gambiae* sufficed to establish both a fully susceptible and a refractory strain that blocks Plasmodium development by encapsulation and melanization of oocysts in the midgut (COLLINS *et al.* 1986). Similarly, a natural refractory mechanism of *A. gambiae*, which lyses ookinetes of the incompatible avian malaria parasite, *Plasmodium gallinaceum*, within the midgut epithelium, can be overcome by genetic selection in a manner suggesting that refractoriness/susceptibility of the mosquito is determined by a small number of loci (VERNICK *et al.* 1995). Field application of the population replacement strategy is likely to depend on mosquitoes that have been engineered through germ-line transformation to be refractory by a well-controlled mechanism. Nevertheless, characterization of natural refractoriness mechanisms undoubtedly will be an important step in

that direction and will be facilitated by a combination of genetic and molecular biological/biochemical approaches. A genetic map is a prerequisite for localization, positional cloning and subsequent characterization of endogenous genes controlling refractoriness. Moreover, such a map is necessary for broadly based genetic studies of anopheline biology, which must underlie any reasonable malaria control strategy that takes account of the insect vector.

In the avian malaria vector *Aedes aegypti*, susceptibility to *P. gallinaceum* is determined mostly by a dominant allele of a second chromosome locus (KILAMA and CRAIG 1969; THATHY *et al.* 1994). Recently, a genetic linkage map of *Ae. aegypti* was developed based on restriction fragment length polymorphisms or RFLPs (SEVERSON *et al.* 1993), and a single locus, *pgs*[2, LF98], was found to contribute >50% of the observed refractory phenotype. An additional locus, *pgs*[3, MaII] was shown to contribute ~10% of the phenotypic variation (SEVERSON *et al.* 1995). Refractoriness of *Ae. aegypti* to the lymphatic filarioid nematode, *Brugia malayi*, is associated with a major and a minor locus on the first and the second chromosomes, respectively (MACDONALD 1962; SEVERSON *et al.* 1994).

The genetics of *A. gambiae*, and specifically the analysis of refractoriness, is less advanced. A serendipitous finding was that, in the laboratory strains used, encapsulation of *P. cynomolgi* B oocysts was associated with an esterase phenotype, EstA/A, and susceptibility with EstC/C (VERNICK and COLLINS 1989). The esterase phe-

Corresponding author: Fotis C. Kafatos, EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany. E-mail: kafatos@embl-heidelberg.de

notype was found to be determined by two separate genes, both highly correlated with a chromosomal inversion (*2La*) spanning chromosome divisions 23–26 (CREWS-OYEN *et al.* 1993), or ~10% of the genome (*ca.* 25 megabases of genomic DNA). A different locus was implicated in the encapsulation of *P. cynomolgi* Ceylon; this locus assorted independently of the esterase phenotype and the chromosomal inversion (COLLINS *et al.* 1986; VERNICK *et al.* 1989). The susceptible response to *P. cynomolgi* B oocysts was shown to be mimicked by failure to encapsulate intrathoracically injected negatively charged C-25 Sephadex beads (PASKEWITZ and RIEHLE 1994).

The limited genetic information on *A. gambiae* is due partly to unfavorable traits of this species, such as its swarm mating habit (MARCHAND 1984). Nevertheless, a few morphological mutations have been described and some linkage studies performed with this species (MASON and DAVIDSON 1966). Modern molecular markers should considerably facilitate genetic analysis. RFLP markers have been identified recently (ROMANS *et al.* 1991). We have focused on genetic markers identifiable by the polymerase chain reactions (PCR, SAIKI *et al.* 1988), so that the relatively few progeny of labor-intensive crosses can be analyzed exhaustively for a large number of markers, each requiring a very small amount of DNA for dependable genotyping. We have reported previously a pilot project in which we used 24 microsatellite markers and a visible white eye marker to construct a map of the *X* chromosome, at an average 2 cM resolution (ZHENG *et al.* 1993). Here we extend this study to the entire genome to an average 1.6 cM resolution using 131 microsatellite markers as well as five morphological markers and one biochemical selectable trait. In addition, we localize 47 of the microsatellite markers cytogenetically, either by *in situ* hybridization to nurse cell polytene chromosomes or by hybridization to chromosomal division-specific DNA pools (ZHENG *et al.* 1991). Thus, an integrated genetic and cytogenetic map that covers all three chromosomes is now available and should greatly facilitate molecular genetic studies on *A. gambiae*.

MATERIALS AND METHODS

Mosquito strains: Strain LU, which is homozygous recessive for the *red-eye* (*r*), *lunate* (*lu*) and dieldrin sensitivity (*Dl^s*) loci, was derived from a red-eye (RE) strain (BEARD *et al.* 1995). Strain DL, which is homozygous for the dominant *r⁺*, *ln⁺*, and *Dl^R* alleles was derived from a white eye (WE) strain (ZHENG *et al.* 1993; BEARD *et al.* 1995).

The dieldrin resistance mutation was identified by selection of adult mosquitoes with insecticide over several generations (W. G. BROGDON, personal communication). The resistance locus selected is possibly the one described by DAVIDSON (1956) and CURTIS *et al.* (1976), although identity cannot be conclusively determined. The *red-eye* mutation is a spontaneous recessive mutation affecting eye color (BEARD *et al.* 1995). The lunate phenotype is due to a spontaneous, recessive mutation that was discovered in the RE strain (M. Q. BENEDICT, unpublished data). The mutation affects the curvature of the lateral setae of the first three abdominal segments in fourth instar larvae. In wild-type larvae, these setae curl forward after eclosion then straighten several hours later, but in mutant larvae they fail to straighten after eclosion.

Genetic crosses: Seven heterozygous F_1 males generated from a single forced pair mating between LU and DL were individually backcrossed, each with one LU female, generating seven families (M2-1, -2, -3, -6, -7, -8, and -9) with 93, 57, 42, 101, 108, 97, and 102 progeny, respectively. All of these families were used to map the morphological markers, but only M2-2 and M2-7 were used for microsatellite mapping in the present study. The eye color and lunate phenotypes of each group of F_2 progeny were scored at the early fourth instar larval stage. Dieldrin resistance and susceptibility were determined by exposing larvae to 1 ppm dieldrin (Chem Service, W. Chester, PA) in water for 1 hr.

We also analyzed five additional families (A–E), which have been described previously (ZHENG *et al.* 1993), from backcrosses between individual SUA/WE F_1 females and WE males. Six additional families, also generated from single pair matings, will be described elsewhere. Four of these families, E2–E5, were derived from crosses between strains L35 and 4arr of *A. gambiae* maintained at CDC, Atlanta, GA (F. H. COLLINS, unpublished results). E2–E5 were informative for mapping of *pink-eye* (*p*, see below) and two (E4 and E5) were also marked with *collarless* (*c*, MASON and DAVIDSON 1966). Families L3 and L4, not marked with any visible marker, were derived from crosses involving SUA and G3 (VERNICK *et al.* 1995).

The available laboratory strains are not extensively inbred and show some intrastrain microsatellite polymorphisms. Therefore, in some cases it was possible to score recombinations in the testcross parent as well as in the F_1 parent (see Table 6).

Microsatellite markers: Phage M13 clones with DNA inserts containing microsatellite sequences were identified, from either size-selected libraries of total genomic DNA or from division 3, 7, 12, 23, or 25 chromosomal DNA pools (ZHENG *et al.* 1991), using as probe ^{32}P end-labeled oligonucleotides, $d(GT)_{15}$ or $d(GA)_{15}$, as described (ZHENG *et al.* 1993). The DNA sequence for each clone was determined and a pair of primers flanking the microsatellite sequence was designed for PCR. Oligonucleotides were synthesized in a Millipore Milligen cyclone plus DNA synthesizer and resuspended at 20 μM.

Genotyping and map construction: Genomic DNA was prepared from each individual larva or adult using a protocol for single adult mosquitoes (ZHENG *et al.* 1993) and was resuspended at ~5 ng/μl. Genotyping was performed as described (ZHENG *et al.* 1993), on a total of 13 families: families A–E (ZHENG *et al.*, 1993), M2-2 and M2-7, E2–E5, and L3–L4. The genotype data were analyzed with the computer program Mapmaker 3.0 (LANDER *et al.* 1987; PATERSON *et al.* 1988), and a composite map was constructed with Joinmap (STAM, 1993) with the linkage LOD score set at 3.0 with a Kosambi (1944) mapping function, which assumes positive interference in crossing over events.

Polytene chromosome *in situ* hybridization: PCR amplifications were performed with microsatellite primers using the corresponding single-strand M13 clone DNAs as templates. PCR products were labeled for use in polytene chromosome *in situ* hybridization as described previously (ZHENG *et al.* 1993; KUMAR and COLLINS 1994); for some markers the entire M13 clone (vector plus insert) was labeled and used as probe. When multiple signals were obtained under the described conditions, lower concentrations of probe DNA and increasingly higher stringencies of washes were used until a unique signal was obtained. Polytene chromosome squashes were prepared from half-gravid ovaries of the mass-reared strain G3,

which is known to carry some inversion polymorphisms. The chromosomal map utilized was a revised version of a previously published map (COLUZZI and SABATINI 1967), made available by the authors (M. COLUZZI, A. SABATINI, M. A. DI-DECO and V. PETRERCA, unpublished data); a derived photomap (A. J. CORNEL and F. H. COLLINS, unpublished observations) was also used.

RESULTS

Genetic markers: Microsatellite markers were identified by screening genomic libraries (cloned in M13 vectors) with labeled oligonucleotides, d(GT)$_{15}$ and d(GA)$_{15}$, as probes for simple sequence repeats. Sequence analysis of positive clones confirmed the presence of tandem dinucleotide repeat arrays, and unique sequences flanking each array were used to design a PCR primer pair that defines the microsatellite marker. A total of 165 dinucleotide repeat markers were obtained by this procedure; one additional marker each was identified from cloned genes of the trypsin family (MÜLLER *et al.* 1993), a homologue of the yeast translation factor *Sui1* (BESANSKY *et al.* 1994), the eye color marker *white* (BESANSKY *et al.* 1995), and an anonymous cDNA clone. One tri- and one tetranucleotide repeat marker were identified from random genomic and cDNA clones, respectively.

Of these 171 markers, 23 could not be scored for various technical reasons; their use would require redesigned primers or recloning (for those that have the repeat array too close to the end of the available clones). Of the remaining 148 microsatellites, 131 (89%) were found to be polymorphic in at least one of the 13 families tested (*e.g.*, Figure 1). Tables 1–3 list these markers in the order of their respective genetic map coordinates (see below). Seventeen markers (11% ; Table 4), including the one from the trypsin gene family locus, could not be mapped genetically because they were not polymorphic in the families tested (data not shown). The simple sequence arrays were interrupted with variant sequences in 30 of the 131 polymorphic microsatellites (23%; Tables 1–3) and in 8 of the 17 nonpolymorphic cases (47%; Table 4).

Conforming to a previously established convention (ZHENG *et al.* 1993), X-linked markers are formally designated by numbers preceded by the prefix *AGXH* (*Anopheles gambiae X* Harvard), second chromosome markers *AG2H*, and third chromosome markers *AG3H*. Markers that by sequence are located in or very near genetically or biochemically known genes are identified by a standard code followed by the symbol or the name of the gene (*e.g.*, *AGXH711/w*).

Genetic crosses: Single pair mating was used to generate 18 families, of which 16 were obtained by crossing two visibly marked strains and backcrossing the resulting F$_1$ individuals with one of the parental strains. Families L3 and L4 were also derived from backcrosses, but were not morphologically marked. Five of the families were used to map only morphological traits, and 13 were used also for microsatellite mapping.

Five families (A–E, ZHENG *et al.* 1993) were marked with an X-linked mutation that results in white eye phenotype and is now designated *pink-eye*$^{\text{white}}$ (*p^w*; see below). Families E2–E5 were marked with a different allele of the same X-linked mutation, *pink-eye* (*p*, BEARD *et al.* 1995; BENEDICT *et al.* 1995). In families M2-1–M2-9, one of the autosomes was marked with the insecticide dieldrin resistance marker *Dl* and the visible marker *lunate* (*lu*). *In situ* hybridization of linked microsatellite markers to the polytene chromosomes (see below) showed that this linkage group corresponds to the cytologically defined second chromosome. The *collarless* (*c*) mutation present in families E4 and E5 was also found to be associated with this chromosome. The remaining linkage group, including the *red* eye (*r*) marker that is carried in families M2-1–M2-9, was identified with the third chromosome by *in situ* hybridization of linked microsatellite markers. Table 5 shows the frequencies of phenotypes in seven backcross families (M2-1–M2-9), which demonstrate the linkage of *lu* with *Dl* and the independent assortment of *r*. These linkage relationships are also shown in Figure 1, together with data that assign two microsatellite markers to the *red-eye* marked third chromosome.

Genetic mapping procedures: A total of 679 progeny (13–108 progeny from each of 13 families) and backcross parents were genotyped for various sets of microsatellites (Table 6). No evidence of mutations occurring during the crosses was seen. Microsatellite sequences with as few as 6 d(GT) repeats were found to be polymorphic, and size differences between alleles ranged from 2 to 60 bp (ZHENG *et al.* 1993).

In anopheline mosquitoes, recombination occurs in males as well as females (see CLEMENTS 1992; and Figure 1). Because some of the autosomal markers were polymorphic in both the SUA/WE heterozygous females and in WE testcross male parents (Table 6), it was possible to make comparisons of recombination frequencies between males and females in families A and B. Recombination frequencies in both sexes were comparable. For example, in family A the distance between *AG3H776* and *AG3H249* was 12.5 cM in the female and 13.6 cM in the male parent.

Initially, the genotype data for each family were analyzed separately, using the Mapmaker 3.0 program (LANDER *et al.* 1987; PATERSON *et al.* 1988). Three linkage groups were unambiguously defined, and consistent maps were obtained in different families in which sufficient markers were scored (A–E, M2-2 and M2-7; Table 6). Therefore, data from all families were pooled and analyzed with Joinmap(STAM, 1993) to construct the final map (Figure 2 and Tables 1–3). Details of the mapping experiments and features of individual chromosome are discussed below.

X chromosome: A pilot map of the *X* chromosome, based on 24 microsatellite markers, has been described previously (ZHENG *et al.* 1993; KNUDSON *et al.* 1995). That map was anchored on a white eye phenotypic trait,
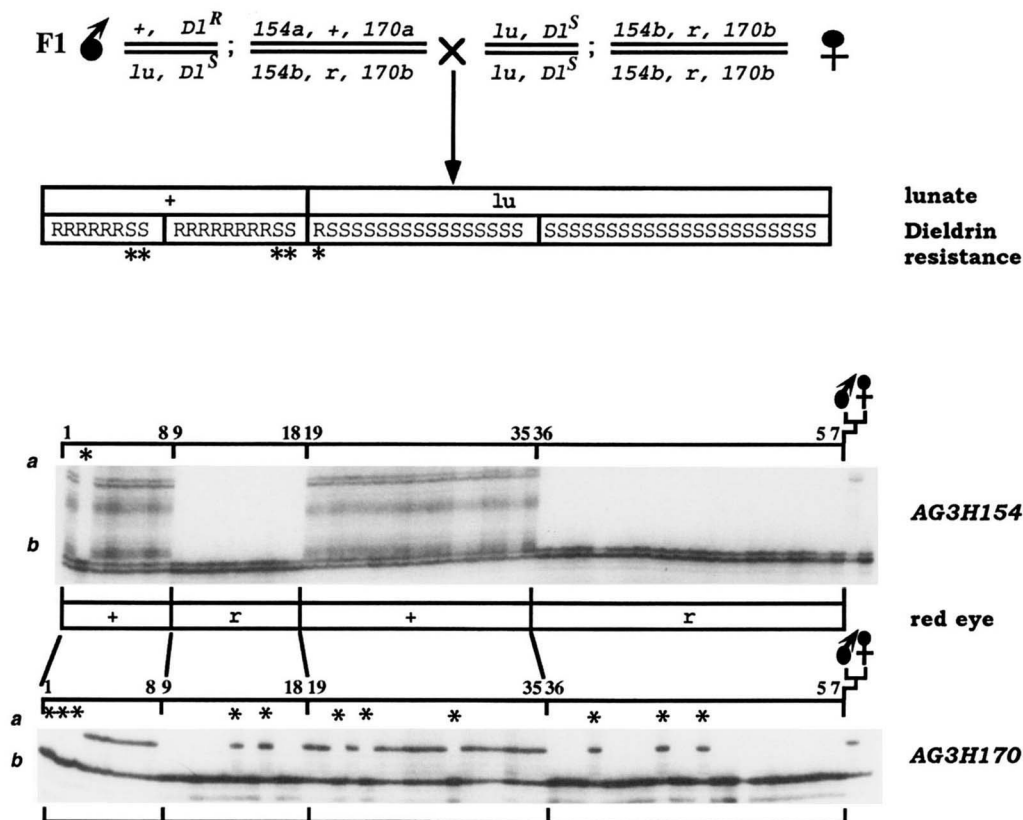
FIGURE 1.—Genetic mapping of morphological, biochemical and microsatellite markers. A DL/LU heterozygous F_1 male was backcrossed to an LU female, generating 57 progeny (family M2-7). The phenotypes of each individual for lunate (+ or lu), dieldrine resistance or susceptibility (R or S), and red eye (+ or r) traits are shown in boxes. Evidently *lu* and *Dl* are linked, showing only five recombinants in 57 progeny (*). The *red-eye* is unlinked to *lu* and *Dl*, in that the *r* and + alleles of the *red-eye* locus segregate at comparable frequencies with the mutant and wild-type alleles of the other two loci. The genotypes for each individual at microsatellite loci *AG3H154* and *AG3H170* were scored by the indicated PCR typing patterns. It can be deduced from the relative frequencies that in the F_1 male parent, *AG3H154b* and *AG3H170b* are linked with the *r* allele, while the microsatellite *a* alleles are associated with the wild-type allele. The exceptional recombinant individuals (1/57 in the case of *AG3H154* and 11/57 in the case of *AG3H170*, relative to *r*) are indicated (*).

which at the time was ascribed to the homologue of the *white* gene of *Drosophila melanogaster*. However, subsequent work revealed that in fact this mutation corresponds to a different gene and is allelic to a mutation known as *p* (BENEDICT *et al.* 1995; BESANSKY *et al.* 1995); henceforth this white eye mutation will be referred to as p^w. A mutation at a separate, nearby locus also gives rise to a white eye phenotype and appears to be the true homologue of the *white* gene of Drosophila (BESANSKY *et al.* 1995). This second gene, *white* (*w*), has been cloned recently and shows no apparent lesion in the p^w strain (BESANSKY *et al.* 1995), confirming the existence of two distinct loci affecting eye color. By serendipity, the *w* gene includes a polymorphic microsatellite sequence (*AGXH711/w*) in the 3′ untranslated region (BESANSKY *et al.* 1995) and therefore can be mapped easily relative to p^w without phenotypic interference (0.4 cM, or 1 recombinant out of 248; somewhat higher values were independently estimated by BENEDICT *et al.* (1995).

In the present study, the average resolution of the *X* chromosome map was doubled by the addition of 22 new microsatellite markers, which were mapped in all five previously used families (A–E). Four additional families (E2-5) were also used. Data from all nine families and all 46 markers were analyzed with Joinmap, yielding a map that spans 48.9 cM and has an average resolution of 1.1 cM (Figure 2).

Second chromosome: The second linkage group was mapped in 11 families (all those listed in Table 6, except L3 and L4). It encompasses 60 markers, spanning 72.1 cM (average resolution 1.2 cM), and was identified with polytene chromosome *2* by *in situ* hybridization of some of the microsatellite markers (see Table 2). As noted above, this linkage group also includes the *lu*, *c*, and *Dl* genes. From the composite map, it appears that *lu* is very close to *AG2H786* (2.1 cM apart); *Dl* to *AG2H772* (1.3 cM apart), and *c* to *AG2H79* (0.4 cM apart).

Third chromosome: For unknown reasons, randomly selected dinucleotide repeat arrays were less frequently cloned from the third chromosome, as compared to the second and the *X*. The map of the third chromosome was constructed using all 13 families listed in Table 6. It encompasses 29 markers, spanning 93.7 cM (average resolution 3.2 cM). As noted above, it also includes the visible marker *r*, which was mapped within a 3.5 cM interval, between *AG3H154* and *AG3H127*.

TABLE 1

X chromosome markers

Marker	Cytol.	Dist.	5′ primer (5′ → 3′)	Repeats	3′ primer (5′ → 3′)	Allele size	Ref.
293		0	ACATCTTTCAGCACCACTGG	(GA) 8	CCGGCGACCACTTGGAACC	91	
145		1.9	TGGTGGAATGTGAGACACAG	(GT) 11	ATGATGGTCGATCCTTGTCC	75	1
77	4B	8.3	TGGGACTGTAAGTGTCTCCC	(GT) 10	TATCAGTGAGGCCGAGTTGC	89	1
503	(4)	11.8	AGGTTAGAGTGAGCAACCAC	(GT) 30	GCACTGCATCTCTCCAATAC	118	1
36	4B	13.1	CGTATGTTTGCTAGGGGTGG	(GT) 14	GTCAAGAAATGGGCCACAGG	196	
484	(4)	18.6	TGCAACCACTACTGGTACTC	(GT) 6 + 4	GATCTCTTGCCCTTACAC	56	
495	(4)	18.7	GCAACCACTACTGGTACTCG	(GT) 8 + 4	GATCTCTTGCCCTTACAC	63	
71		21.4	GCGGAGTTATTTCCTGAACC	(GT) 8	ACAGGCCAAGCAAATGCAGG	141	1
755		21.5	ACACCTCCATGAACGGTTCG	(GT) 9	ACAGAACAGACAAGGCAGCG	121	
53	3D	25.0	GTTTCGGGGCTTGAGAAGTG	(GT) 7	CTTCACGTGGCTTTGCTGTG	96	
540		25.9	GCGAAATGGAACACGAGCTT	(CA) 11	GTCGATTGGCGTTCATCTTG	114	
106		26.1	CTCTTGGCTTACGCTCCTGG	(GT) 4 + 12 + 8	GGGAATGAAGATGAGAAGCC	120	1
180	3C	26.6	GTATGTTGTGATCTCCTGCC	(GT) 10	AAAACGAGCCACCACCAGAG	72	1
289		27.0	CTGCGAACTTTGCTGATTCG	(GA) 9	TTCGCCAAACTGACAACTGC	90	1
465	(3)	27.4	GCGACTACTTGTGTTTGAGC	(GT) 12	TTTTCTACACTCCGGGAGCG	85	
808		27.5	TCCGCAAAACAACCGATGAC	(TG) 7 + 5	CAGCAGGACAATCACTACGG	118	
471	(3)	29.1	TAGCCAACCGAAGAAGTGAC	(GT) 11	CGGTGTCTGCTTGTCTTCTG	75	
19	(3)	29.8	CTTTTTCTCCCCATTATCTC	(GT) 9	CTGCAGTGTCCATTACGTAC	71	1
38		31.2	AGTGACTACGCTTCTCGGAG	(GT) 7 + 4	AAGTCATACTCTTGCGCCCG	98	1
768		33.3	GAAAGTAATGATCGTCCCGC	(GT) 11	TTAACATTTTTGGCCACCCC	112	
1002	3C	34.9	GATCGGTATATGCTTCCCGC	(GA) 23	AATAAGCCACGGCGTATCCC	149	1
179		34.9	CCATCCCCTCGACAGACC	(GT) 10 + 6	AACGACGTAAGCTGACACGG	106	1
454	(3)	35.3	GATCTTCACCACCGGAGAG	(GA) 8 + 6	CTTTTCTTTGCTCTCGTCGG	95	
131		36.1	TTCCCACACTTTCTCCCAGG	(GT) 48	ATAATGCGCTGCTCCCAAGG	175	1
459	(3)	36.5	GGTGCTTTTCTACCTATGGG	(GT) 10	ATCGCGCTGAAATTTCCCGG	96	
784		37.4	TGGTGAAAGAACAGACCCCG	(GT) 11	TGTAACGGGCAAGAAAAGC	109	
472	(3)	37.4	CATCTCCATGCGCTACGAAG	(GT) 29	ATCCGTACCTCCACCACAAA	130	
24	(3)	37.7	GGAGGCTAAAATCACGGTTG	(GT) 36	GATCGGCAAGACTATCGGCC	132	1
25	3A	37.7	GCCGAAAACATTCCAACAGG	(GT) 9	CAGTTATGTCGGCATGCTAC	134	1
80	2B	37.7	TGCTCTCTCCTACATCGAGG	(GT) 9	GCCAGTGCTCTAGATTAACG	81	1
805		38.3	TGTTGCGCCTGAAAGGTAGG	(TG) 4 + 12	CGAGATGCTCTCTTCGATGG	124	
99	2C	39.5	CGGGAATTTGTTGCTTCCTG	(GT) 8	TCGCCCTCTTTCTCCATCTC	127	1
p		40.4					
711/w	2A	40.8	CCCACAGCAAAACGAGAATG	(GA) 9 + 4	GACAACTTGCATTTCACTATG	131	2
49	1D	40.8	CAGCGCCTCCATATAGAACG	(GT) 5 + 4	GATCATTCAGCTGAACCTGC	97	1
7	1C	41.4	CACGATGGTTTTCGGTGTGG	(GT) 8	ATTTGAGCTCTCCCGGGTG	99	1
810		41.7	CGTCTCGAGAGATCGATAGC	(GT) 7 + 8 + 7	TGATGGCCTTAACATGGTGG	123	
261		41.8	GGAGGCATTCTTTTCATTCC	(GT) 3 + 4 + 4	TTCGCTTTCCCAGCGGGTGG	117	
81		41.9	CACTGTAAATCGGAAGCGCG	(GA) 7	CGGGCGGTTAAAGAAAACGG	106	1
32		42.1	CGGTGCGTGTTCCTCGTGC	(GT) 29	TATGGTGTGGTTTCCCGTCC	120	1
8	(1)	42.5	GGATGTGCTCCCAATACAAG	(GT) 4 + 6	CTTATCGCACTGCAAGTGTC	127	1
253		43.0	TGCTGGTTCTGGCTTTGTGG	(GT) 18	GTCTGTCGCAATCGGCTTCC	116	1
37	1C	43.3	ATGTCTTGCTCACCTCGAGC	(GA) 9	TAAGTTGGGCGTCTTGCTGG	109	
766		43.3	CAGGTAGTAGGAGTAGATGC	(TG) 5 + 6	AATTATGAGCACGGTGGGTG	95	
100		43.9	AGAAAGGAAATGTAACGCGG	(GA) 7	CTTTCATCTTGGCTGGCTGC	93	1
412	(6)	47.9	GCATGCACCTGTTGGGACAG	(GT) 16 + 4	AAACCTTACCCAAAACACAG	179	1
678		48.9	CCTCTCCCCAGAATCGGTAC	(AG) 7	AAGAGCAGAAACAACCGCAG	153	

All *X* chromosome microsatellite markers have a prefix of *AGXH* followed by the number indicated in the first column. The *p* indicates the *pink-eye* morphological marker and *711/w* a microsatellite marker in the *white (w)* gene. The second column (Cytol.) shows the cytological locations in the polytene chromosomes, if known; the numbers in parentheses indicate that the marker was identified from a division-specific DNA library, other numbers and capital letters indicate localization to chromosomal divisions and subdivisions by *in situ* hybridization. Note that recent experiments showed that markers *AGXH25* and *AGXH180* were located at 3A and 3C, rather than 3D and 3B, respectively, as reported previously (ZHENG *et al.* 1993). Cumulative genetic distances (Dist.; in *centimorgans*) from the most distal markers are shown in column 3. The PCR primer pairs are shown in columns 4 and 6, flanking the number of simple sequence repeats in the cloned allele (shown in column 5); + sign in column 5 indicates interrupted dinucleotide repeats. In column 7, the overall sizes of the cloned alleles are shown in base pairs. Column 8 (Ref.) shows references: 1, ZHENG *et al.* (1993); 2, BESANSKY *et al.* (1995). Note that *A. gambiae* is fixed for a chromosomal inversion *Xag* (COLUZZI *et al.* 1979), so that the chromosomal division beginning at the telomere is 4; centromere is at 6.

***In situ* hybridization:** The three linkage groups were anchored to the polytene chromosomes by *in situ* hybridization using cloned microsatellite markers as probes (Figure 3). Some of the markers with very short unique sequences and long dinucleotide repeats could not be localized unambiguously. Tables 1–3 and Figure 2 show that, in general, the cytogenetic order of the mapped microsatellite markers is consistent with their genetic order. One exception is the cytogenetic order of the microsatellite markers around *2La* (divisions 23–

TABLE 2

Second chromosome markers

Marker	Cytol.	Dist.	5' primer (5' → 3')	Repeats	3' primer (5' → 3')	Allele size	Ref.
417	(7)	0.0	GTGCAGGAGATGTTCCTACC	(GT) 9	GATATGGTGTTTAGGCCCAG	89	
794		0.3	ACGATAAAACGCGGATTGCG	(GT) 4 + 9	CCGTGGACGGGTTTCAAACG	97	
46	7A	0.7	CGCCCATAGACAACGAAAGG	(GT) 8	TGTACAGCTGCAGAACGAGC	138	1
784		2.1	TGGTGAAAGAACAGACCCCG	(GT) 11	TGTAACGGGCAAGAAAAGGC	109	
793		2.5	CCTTTATCGGAGAGCCTGAG	(GT) 13	ATTAGAGCTAACGGGCACGC	102	
175		3.1	AGGAGCTGCATAATTCACGC	(CA) 8	AGAAGCATTGCCCGCATTCC	97	1
442	(7)	3.7	GTGCGATGAGGTGTTTAGGC	(TG) 9	GTGCGATGAGGTGTTTAGGC	93	
290		4.3	GGCGTGTGCTGTGCTCCC	(GT) 9	ACGCAATTTTTGCCTCAGCG	98	
427	(7)	6.0	TCGGCAGTGAAAGAATCCGC	(GT) 10	CCTCTACTCAGCACGAATGG	131	
803	7B	8.0	CTCGATAAATCCCGTCGGTG	(TG) 11	GTCGGTTTGAGGTTGTAAAGC	112	
157	7C	11.5	ATCCTGCTCCCCATAAAGCC	(GT) 13	CGTGCTCCAAGGTCTACTTG	91	
815		12.4	GTGATATTTATGGGCAACTGC	(GT) 19	TCAGAGTGTGCTATCCTTCG	98	
197		13.4	TACCTCTGTGTTCGGTTTCC	(GT) 8	GGTGGTATGGCGATGGAACG	85	
788		14.3	TGGTGTAGAGCATCGTTACC	(GT) 8	GGTATACCAGTGAGTTTCGC	85	
156		17.8	AGCCTGTCAGAAATCTTTCC	(GT) 11	CCCTCTTCCCAAAACACCAC	78	
791		18.6	TCGAGCTGCTCTTCTTAACC	(GT) 18	CGCTTCCTTTCAAAACATGC	116	
799	9B	18.7	TTATGGGCAACTGCGGATGG	(GT) 18	CGTGCGTTTGATACATCTACG	95	
769		20.4	TTATGGGCAACTGCGGATGG	(TG) 19	AGTGTGCTATCCTTCGTGCG	108	
79		24.3	CGGGTAGCGCTAGAAGTATG	(GT) 20	AGAGAAATGTGCCGAAGGGG	201	
c		24.7					
187		24.7	CCGGAGCAGAGATAAACAGC	(GT) 48	CACAGACGTACACCTAATGC	155	
85	11A	25.9	ATTTATCATACGGCGCCCAC	(GT) 11	TTGAAAGGTTGCAACGAGCGCG	153	
522		27.3	GGTGTGTTTTTTTTTCCTCGC	(GT) 12	CGTGAAGTGTAAAACGGAAGG	108	
1010	11B	27.7	GCGTATGTCAATGGCGAGAA	(GATA) 6	TCGCTGGAAATTGTCACACC	117	
57		27.8	GATCATTTGCTTAACCAAAC	(GT) 12	GGGGACCACAATGGAAAGTC	129	
1003		28.4	GCTTAACCAAACATTTTTTCCC	(TG) 12	GGACCACAATGGAAAGTCGG	118	
lu		30.8					
786		32.9	TGTGAAGCATTTCCTTGGCG	(GT) 11	TGCCCTTGAGTCGAGGTAGC	85	
757		33.0	TGATCGCGCCCAATCAATCC	(GT) 6	ATCGATCGTACAGATGTGCC	94	
26	(12)	34.1	GGTTCCTGTTACTTCCTGCC	(GT) 8 + 29 + 4	CCGGCAACACAAACAATCGG	154	
590		37.8	CGGGAAAGCGAAGTGTACGA	(GT) 11 + 8	TGCGGCTGGTGAACATTTTC	125	
102		39.3	TTGTGGGAGCTGCTGGTGC	(GCT) 7	TATGGATCAGCGCACCACTG	108	
125		40.4	AGGAGCATAACACATCGCCC	(GT) 11	CGCTCGTCAAAGAAACTGGC	102	
95	13B	42.3	CCGGGGTCAGTGTCAGTGG	(GT) 5 + 2 + 2	CCGTTTGCGTCCAATTTTCC	110	
720		43.5	ATTAGAATCCAAACCAGCGG	(GA) 8	ATAAGCTAATGCGCTGCTCC	69	
60	14B	45.3	TGTTTGGGACGGAACCGAG	(GT) 10	TCTCGTGACGGATGATACTG	166	
801		46.9	AAATGCCATCAACGCGAAGG	(TG) 18	GTGGATCACCATTCGATGAG	108	
135		47.3	TCATGCACTGTTTGCTCGGC	(GT) 7	CTGCCCCATTCAATTGCAGC	103	
778		48.0	CGCGTTAAAACGATGAAGCG	(GT) 9	TGAAGACACCTCTTGCGTGG	129	
796		48.7	CTTTGCCATTGCACGGTCCC	(GT) 10	TTCGGCTCCGCTCACTCAAC	89	
770		50.1	CAACATGGAGGCGCATGATC	(GT) 8 + 4	GCGTTCCATCGAAATCAGAC	164	
147		50.8	CTGCTGTTGCTGCCAAAATG	(GT) 8	AGCTTCACGGAAAGCAAAGC	177	
Dl		51.9					
772		52.9	TACAGCTGTTTGGGAGTTGG	(GT) 8	GGGTCGGCTTTTATTTCCTCG	116	
11/Sui1	24D	53.8	GGCGAGCAGTTCATTCAAGT	(CT) 10	CGTCTGGAAGTTTCGTTGAG	125	2
215		54.1	GCAACTTGTTGGCTGTTAGC	(GT) 12	TCTGCAAGCGTTAAGCACTC	80	
603	(25)	54.7	TGCACCGTTGATGCACATGC	(GT) 8	GTGGACGATGTGAAAGATAAGG	109	
164		54.9	GTGGTACCTCTGTCATACCC	(GT) 10	ACAACAAAAGGCACCGCAGC	92	
161		55.0	TTCACCTGTCCCGTGTGGTC	(GT) 8	GGAACTTTCGGTGCTGTAGG	92	
819		55.1	ACGATAGCACGTTTCTGTCG	(GT) 6 + 3	TTATGCTTTGCCGAATCGCG	116	
523		55.2	CTCGTTAGGCGCTTGTGAAC	(GT) 19	CACTTCACGACTGTGAGCAC	188	
143		56.3	CGTACGAGTGAGTGAGTTGG	(TC) 9	CAAAAATAGCATCACGGCCG	160	
802	26D	57.0	TTTGGTGAGGGGTTTGTTCC	(TG) 31 + 24	GGCAATTCATTCGTCACACC	227	
1		58.6	CTTTTACACCGAGGGAAAG	(GT) 18 + 70	CGACCGTACACATAAACAC	214	
637	(23)	60.3	TCGAAATGTATGCGAAATGCAG	(CA) 5 + 6	CCTTCTTTCCTCGATGCATTCC	107	1
787		61.9	CGGGTCAAAGAAAACTCAGC	(GT) 3 + 12	GCATAAGAACGGCACATTGC	124	
117		65.1	CGGAACGCACGGAACAATTG	(GT) 6 + 4	CGTTGCAGATTTCCCAAACG	100	
1012	23C	68.5	AGTGTTCAGAGCGGGAAAAG	(GC/GT) 6 + 9	GTACAACCCGCAGGAGAAAC	131	
675		70.4	CGTGACACTTTCAGGACACC	(AC) 10	GGCAAAAGGCTGGAAAACCG	122	
325		72.4	CCGGTGTCCGTGTTG	(GT) 13	GCGCGAAAGCAAATGACACG	100	

Same conventions as in Table 1, except that the microsatellite prefix is *AG2H*. The *c* (*collarless*) and *lu* (*lunate*) markers are exclusively morphological; *Dl* (*dieldrin resistance*) is a selectable marker, and *11/Sui1* is a microsatellite marker in a cloned homologue of the yeast translation factor, *Sui1*. References: 1, Lanzaro *et al.* (1995); 2, Besansky *et al.* (1994). The telomeric divisions for 2R and 2L are 7 and 28, respectively, the centromere is between divisions 19 and 20, and the large polymorphic 2La inversion encompasses divisions 23–26.

TABLE 3

Third chromosome markers

Marker	Cytol.	Dist.	5' primer (5' → 3')	Repeats	3' primer (5' → 3')	Allele size	Ref.
93	29A	0	TCCCCAGCTCACCCTTCAAG	(GT) 4 + 7	GGTTGCATGTTTGGATAGCG	209	1
776		0.9	TGCGGATCATAATCGAGTCC	(GT) 7	TCACAAACACGCAACGAGTC	96	
128		3.0	CGGGACGGCTAGATAAAGCG	(GT) 21	CCGGGCGACATAACCCACCC	117	
746		3.7	TGGGTTCGAAATTCGCCAAC	(GT) 14	GACGTGTGCACCCGTTGTG	105	
59	29D	4.7	CCCCTATTAAACCCTGGACG	(GT) 9	TGTTGTTGCCCTGCGTTACC	123	
812		6.2	CTGGCCCATTTTGCATATGC	(GT) 10	TGCTCCACCCAAACCACATC	131	
249		11.9	ATGTTCCGCACTTCCGACAC	(GT) 15	GCGAGCTACAACAATGGAGC	129	
119	35B	29.1	GGTTGATGCTGAAGAGTGGG	(GT) 6	ATGCCAGCGGATACGATTCG	174	
555		37.1	GCAGAGACACTTTCCGAAAC	(GT) 8	TGTCAACCCACATTTTGCGC	81	
83		40.9	TTTGTTCCAAACGGGGTGCG	(GT) 6	CCAGCGCGAACTATGGGC	70	1
170		45.9	TATACCCGATATCGTGCGCC	(GT) 10	GGAAAAAAACATCCCCCAAAC	108	
158		46.8	CTGGCACGATCAATCAATCG	(GT) 12	ACGATGGTGTACACGTAACG	96	
341		51.4	CCCAAAGCAATGAACCTCGC	(GT) 13	AGTAGAAGAAGAGGGCAGCG	130	
762		58.7	TTCCCAATTGTCATCGTCCG	(GT) 11	GCGGGTTTTTTTTTACTGGC		
311		61.7	CGGCAGAGGCGCTGCGGG	(GT) 9	CCCCGAAAAACGAAACGCAC	83	
88		61.8	TGCGGCGGTAAAGCATCAAC	(GT) 9	CCGGTAACACTGCGCCGAC	176	
811		65.2	AACCCACAGTACAGCTCGCG	(TG) 9	GTTGCTGCATACTAACCTCG	130	
750		68.0	GCAAAAAAGCTTCTCCCC	(GT) 8	TTAGCTACCGTCGACGCTTC	87	
127	39A	68.2	CCTCTAACTCGATTACCGTG	(GT) 12	GTCAGGGAATTGGAAAGAGC	84	
r		70.9					
154		71.7	TGCAACACTGTGTGCGAGTC	(GT) 10	CGCGGGATAAACCATCAACC	114	
577	42A	71.9	TTCAGCTTCAGGTTGGTCTC	(GT) 16	GGGTTTTTTGGCTGCGACTG	113	
544		74.6	GGATGACCCACTTTACCGCT	(GT) 7	CTCGCCCTTTTCCCCTTACC	108	
312		80.0	TAAACCATCAACCAGCCACC	(GT) 10	ACTGTGTGCGAGTCGGTTGG	108	
758		81.0	TGATTTGCCAGTTCTGCCAG	(GT) 11	GTGATTGGAGTGGCTAGTGG	105	
765		88.0	AGGCCAATGAGGTATCGAGC	(GT) 5 + 6	GCATGGCAACCGTTTTCTGC	85	
753		88.7	CGAACACGTCAATTCCTCAC	(GT) 5 + 7	CCATGGCAACCGTTTTCTGC	140	
242		91.3	TTCATTTCCACCGCAGCTGC	(GT) 8	GGCGACACTCAATCCTTCC	69	
817	44B	93.7	ACTGGTCCGTTGCTGCGCG	(GT) 8	ATGAGTGAATGGTGCGCTGG	124	

Same conventions as in Table 1, except that the microsatellite prefix is *AG3H*. The *r* (*red-eye*) marker is exclusively morphological. Reference: 1, LANZARO *et al.* (1995). Note that the telomeric divisions for *3R* and *3L* are 29 and 46, respectively, and the centromere is between divisions 37 and 38.

TABLE 4

Nonpolymorphic microsatellite markers

Marker	Cytol.	5' primer (5' → 3')	Repeats	3' primer (5' → 3')	Allele size	Ref.
AGXH6	(1)	CAAGGCGTTGACACTGTTCG	(GT) 4 + 6	CTTATCGCACTGCAAGTGTC	203	
AGH91		AAGCATCCTGTCACCAAGCC	(GT) 6	GCCACACATTCACAGCGATC	111	
AGH101		ATCGCGGCATGATAGAGTGG	(GT) 5 + 6	TTGTCTCGAGTTTTCTGGCC	87	
AGH141		CGGAGCAAATCTGAACCGTG	(GT) 9	CCTTGGCCACAACAACATCG	116	
AGH150		TGTGCGCCCGATTGTATAGC	(GT) 20	AACATGGCTCCTGCAGAAGC	146	
AGH177		CCGGAATCGATCGTCAGCG	(GT) 3 + 3 + 4	CCTTACCCCTTCTCTTTGCG	168	
AGH199		TGTATCTGTGTGCACTGTCC	(GT) 4 + 4	ATCTTCGGTGGTGGCATCTG	92	
AGH267		GTTTTGCCACGTTACTCGTG	(GT) 7 + 8 + 7	CACAAACCCGCGCACTTGCA	165	
AGXH450	(3)	TTTTCTATCCACCTTCGCCG	(GT) 4 + 6	CCTAGCATTTCGTTTTCCCC	74	
AGXH487	(4)	TTCCATTCGGTAGCAAGCCG	(GT) 10	GGGATTGGGTTAGTTGCAGC	87	
AGH760		TGTCCGTGCTTAAGTCATGC	(TG) 5 + 4 + 4	ACCCATAACAGGTCAAGAGG	97	
AGH761		TAGTGCACAGCTGCCAGTAG	(GT) 8	GGGTGTGCATTTAGGCAGTG	99	
AGH783		CGTGCGTGCGTGTCAGTTC	(GT) 8	TAATCACGCACAAGTGGACG	111	
AGH786		TGTGAAGCATTTCCTTGGCG	(GT) 11	TGCCCTTGAGTCGAGGTAGC	85	
AGH790		AGTCGCAGTGCAGTGTGGTG	(GT) 11	TGCGCTAGTTTTTCAGCAGC	124	
AGH804		ACGTTTGGGAAAGATCAGCG	(TG) 22	AAAGCTTCGTGCAAAAGGGC	118	
AG3H2/ty	30A	TACAGAGGGAGACTAAGAGG	(GA) 8 + 5	ACATTAATGCTCCCTGTGCC	80	1

Same conventions as in Tables 1–3. The *AGH* prefix indicates markers of unknown cytological locations. Reference: 1, MÜLLER *et al.* (1993, 1995).

26), which is not the same as the genetic order (see DISCUSSION). One small discrepancy for X-linked markers that was observed earlier (ZHENG *et al.* 1993) was resolved with the help of a photomap of the polytene chromosomes: *AGXH180* was mapped at 3C rather than 3B, making its position consistent with the genetic map. However, another small discrepancy persists for *AGXH80* and *99*, which may be due to an undetected inversion. On the other hand, the inverted orientation of the division 1–4 block of the X chromosome, relative to divisions 5 and 6, corresponds to the known inversion *Xag*, which is fixed in *A. gambiae* (COLUZZI *et al.* 1979).

DISCUSSION

We present a genetic map of *A. gambiae* consisting of microsatellite, visible and biochemical markers. Three linkage groups have been defined and assigned cytologically by *in situ* hybridization of selected microsatellite sequences to nurse cell polytene chromosomes. The map includes 131 microsatellite markers and covers a total of 215 cM at an average resolution of 1.6 cM. In addition, 17 microsatellite markers could not be mapped genetically due to lack of polymorphism (Table 4), but provide sequence tagged sites (OLSON *et al.* 1989) for future physical mapping.

Recombination was observed in both males and females, with comparable frequency. Other anopheline mosquitoes also exhibit comparable recombination frequencies in both sexes (MITCHELL *et al.* 1993; SEAWRIGHT and NARANG 1993).

The recombination frequency does not correlate closely with physical distance, as judged by polytene chromosome length. We have noted earlier that division 3 and 4 markers span almost two-thirds of the X chromosome genetic map (ZHENG *et al.* 1993; see also Table 1 and Figure 2). Similarly, markers from six out of 22 divisions (7–12) span almost half of the genetic map of chromosome *2* (Table 2 and Figure 2). Low recombination frequencies for markers in *2L* might be

due in part to polymorphism for a chromosome inversion (see below). In *D. melanogaster*, it has been noted that the frequency of recombination per unit length of polytene chromosome is generally maximal near the middle of each chromosomal arm (ASHBURNER 1989), but it is premature to make a generalization concerning the relationship between genetic and physical maps in Anopheles.

Any genetic map is limited by the size of the data sets that have been used to construct it and by undetected polymorphisms in gene arrangement. The map presented here is the statistically established best estimate from the described data sets. Because of the availability of polytene chromosomes, we have been able to subject this map to a stringent test and confirmed its robustness from the overall consistency between the recombinational and cytogenetic order of the markers. Some variation can be expected because the *A. gambiae* complex is notorious for fixed and floating inversions (COLUZZI *et al.* 1979). The commonly available mapping stocks, including the G3 and WE strains used in this study, are neither homosequential nor devoid of intrastrain inversion polymorphisms. Overcoming this limitation would require an extensive program of selecting inbred lines that are chromosomally homosequential, but sequence-divergent, and robust in the laboratory.

A few minor inconsistencies and uncertainties in the order of very closely located markers in the present map could result from undetected chromosomal rearrangements as well as limitations in the data such as genotyping errors and the pooling of information from many families, each analyzed for only partially overlapping subsets of markers. The map shows only one significant inconsistency, the inferred cytogenetic order of divisions 24, 25, 26, 23 in the *2L* arm (proximal to distal, centromere to telomere). These divisions are encompassed in the large inversion, *2La*, for which *A. gambiae* is highly polymorphic. The order in the wild-type karyotype is 23, 24, 25, 26, whereas it is 26 to 23

TABLE 5

Morphological and biochemical phenotypes of the progeny of seven backcross families

Family	+				*lu*				Total
	Dl^R		Dl^S		Dl^R		Dl^S		
	+	r	+	r	+	r	+	r	
M2-1	20	24	0	1	1	3	20	24	93
M2-2	6	8	2	2	1	1	16	21	57
M2-3	8	7	0	0	1	2	13	11	42
M2-6	23	23	2	5	10	4	19	15	101
M2-7	20	24	5	3	9	6	22	19	108
M2-8	26	24	7	8	2	8	14	8	97
M2-9	19	29	1	3	9	4	19	18	102
Total	122	139	17	22	33	28	123	116	600

The progeny were grouped according to their phenotypes. Evidently, *lu* and *Dl* are linked and are 100/600 = 16.7 cM apart. The *r* mutation is unlinked to *lu* and *Dl*. Columns 4–7 represent recombinations between *lu* and *Dl*.

TABLE 6

Genotyping performed for each family

Family	No. of progeny	Markers	
		Polymorphic	Nonpolymorphic
A	63	*1, 7, 8, 19, 24, 25, 26, 32, 36, 38, 46, 49, 53, 57, <u>59</u>, 60, 71, 77, 79, 80, 81, 83, 85, 88, <u>93</u>, 99, 106, 117, 119, 125, 128, 131, 135, <u>143</u>, 145, 147, <u>156</u>, 157, 158, 161, 164, 170, 175, 179, 180, 187, 197, 215, 242, <u>249</u>, 253, 261, 289, 290, 293, 295, 311, 325, 412, <u>417</u>, 427, <u>442</u>, 454, 459, 465, 471, 472, 484, 495, 503, 510, 522, 523, 525, 540, 544, 555, 577, <u>590</u>, 603, 612, <u>637</u>, 675, 678, 718, 720, 746, 750, 755, 757, 758, 762, 766, 768, 769, 770, <u>772</u>, <u>776</u>, 778, <u>784</u>, <u>787</u>, 791, <u>793</u>, <u>796</u>, 799, 801, 802, 803, 805, 808, 810, 811, <u>812</u>, 815, 817, 819, 1002, 1003, 1010, 1012, 11/Suil, 711/w, p^w*	*37, 100, 127, 154, 753, 765, 788, 794*
B	45	*1, 7, 8, 19, 24, 25, <u>26</u>, 32, 36, 37, 38, 46, 49, 53, <u>57</u>, <u>60</u>, 71, 77, 79, 80, 81, 83, <u>85</u>, <u>88</u>, 93, 95, 99, 100, <u>102</u>, 106, <u>117</u>, 119, 127, 128, 131, 135, <u>143</u>, 144, 145, <u>147</u>, 154, <u>156</u>, 157, 161, 170, <u>175</u>, 179, 180, <u>187</u>, 197, 215, 249, 253, 289, 293, 311, 312, <u>325</u>, <u>341</u>, 412, 417, 427, 442, 454, 503, 523, 577, <u>590</u>, 603, <u>637</u>, 678, 746, 753, 755, 765, 766, 768, <u>769</u>, 770, 776, 784, 794, <u>802</u>, 803, 812, 815, 819, 1002, 711/w, p^w*	*750, 788, 794*
C	76	*7, 8, 19, 24, 25, 32, 36, 37, 38, 49, 71, 77, 80, 81, 100, 106, 131, 145, 179, 180, 253, 289, 293, 454, 459, 465, 471, 472, 484, 495, 503, 678, 766, 768, 803, 808, 1002, 711/w, p^w*	*99, 261, 765*
D	13	*7, 8, 19, 24, 25, 32, 36, 37, 38, 49, 59, 71, 77, 80, 81, 99, 100, 106, 131, 145, 157, 179, 180, 253, 289, 293, 454, 459, 465, 471, 472, 484, 495, 503, 540, 590, 678, 766, 768, 784, 805, 808, 1002, 711/w, p^w*	*249, 765*
E	51	*1, 8, 19, 24, 25, <u>26</u>, 32, 36, 38, 46, 49, 57, 71, 77, 80, 81, 99, 100, 106, <u>125</u>, 131, <u>143</u>, 145, 147, 161, 179, 180, 253, 289, 293, 311, 312, 324, 454, 459, 465, 471, 472, 484, 495, 503, 523, 540, 577, 590, 678, 766, 770, 768, <u>769</u>, 778, <u>784</u>, 788, 805, 808, 1002, 1003, 711/w, p^w*	*7, 88, 157, 765*
M2-2	108	*1, 79, 117, 125, 127, 154, 170, 175, 242, 427, 577, 637, 720, 750, 757, 765, <u>769</u>, 772, 776, <u>786</u>, 787, <u>788</u>, 796, 803, 811, 819, 1010, lu, r, Dl*	*46, 60, 88, 147, 296, 459, 465, 471, 472, 484, 495, 758, 770, 778, 784, 805, 808, 812*
M2-7	57	*1, 79, 117, 125, 127, 154, 242, 427, 577, <u>637</u>, 675, 750, <u>757</u>, 769, 772, 776, 786, <u>788</u>, 796, 803, 811, 819, 1010, lu, r, Dl*	*46, 60, 88, 147, 175, 296, 459, 465, 471, 472, 484, 495, 758, 770, 778, 784, 805, 808, 812*
L3	70	*19, 143, 503, 746, 750*	
L4	63	*83, 119, 127, 170, 341, 555, 577, 746, 750, 765, 811*	
E2	36	*60, 77, 99, 106, 603, 772, 802, p*	*637*
E3	36	*60, 77, 99, 106, 603, 772, 802, p*	*637*
E4	36	*26, 60, 77, <u>79</u>, 99, 106, <u>125</u>, 187, 603, 772, 802, p, c*	*19, 93, 215, 523, 637, 675, 1010, 1012, 11/Suil*
E5	25	*<u>26</u>, <u>60</u>, 77, <u>79</u>, 99, 106, <u>125</u>, 187, 772, 802, p, c*	*19, 93, 215, 523, 637, 675, 1010, 1012, 11/Suil*

Only the genetically mapped markers are listed here. Some tested markers were not polymorphic in either parent of one or more of the 13 families (last column). Underlined markers were polymorphic in both parents.
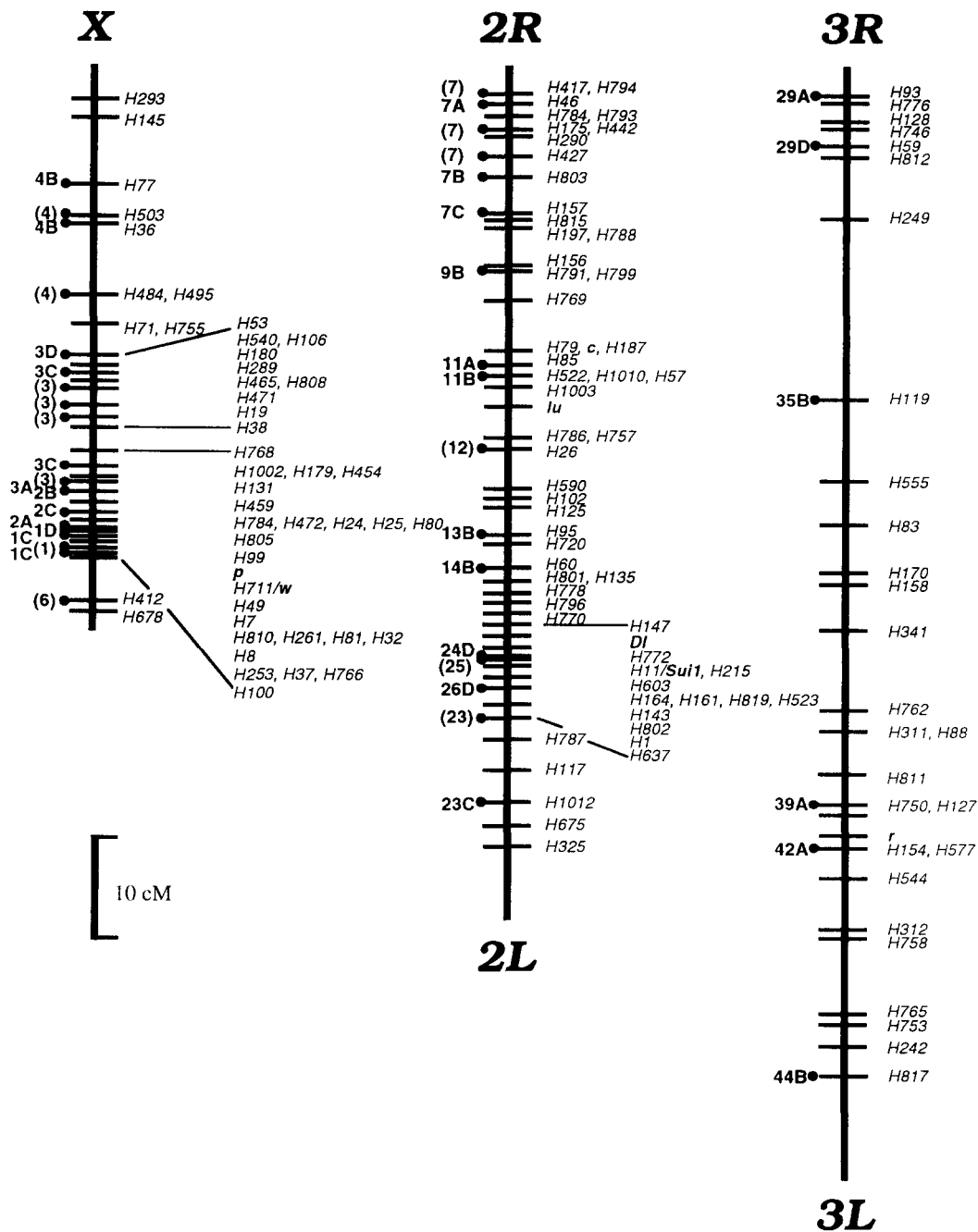
X

H293
H145
4B — H77
(4) — H503
4B — H36
(4) — H484, H495
H71, H755
H53
H540, H106
H180
3D — H289
3C — H465, H808
(3) — H471
(3) — H19
(3) — H38
H768
3C — H1002, H179, H454
3A (3) — H131
2B — H459
2A 2C — H784, H472, H24, H25, H80
1D — H805
1C — H99
1C (1) — p
H711/w
H49
(6) — H412
H678 — H7
H810, H261, H81, H32
H8
H253, H37, H766
H100

10 cM

2R

(7) — H417, H794
7A — H46
H784, H793
(7) — H175, H442
H290
(7) — H427
7B — H803
7C — H157
H815
H197, H788
9B — H156
H791, H799
H769
H79, c, H187
11A — H85
11B — H522, H1010, H57
H1003
Iu
H786, H757
(12) — H26
H590
H102
H125
H95
13B — H720
14B — H60
H801, H135
H778
H796
H770
24D — H147
(25) — Dl
H772
26D — H11/Sui1, H215
H603
(23) — H164, H161, H819, H523
H143
H802
H787 — H1
H637
H117
23C — H1012
H675
H325

2L

3R

29A — H93
H776
H128
H746
29D — H59
H812
H249
35B — H119
H555
H83
H170
H158
H341
H762
H311, H88
H811
39A — H750, H127
42A — r
H154, H577
H544
H312
H758
H765
H753
H242
44B — H817

3L

FIGURE 2.—Genetic map of *Anopheles gambiae*. The map was derived from genotype data collected from 13 families of crosses and analysed with Joinmap (STAM 1993) at linkage LOD of ≥3.0. Each chromosomal arm is labeled at the telomere, but the location of the centromere was not precisely determined (see Tables 1–3). A scale of genetic distance in centimorgans is shown. Microsatellites (with an abbreviated prefix *H*), and visible or biochemical marker loci (in bold italics) are shown on the right. The polytene chromosome positions of the indicated markers are shown on the left side of each chromosome (parentheses indicate origin of the marker in a division-specific DNA library).

in the case of *2La* (COLUZZI *et al.* 1979). It may be pertinent that we detected what appears to be local suppression of recombination due to *2La* inversion heterozygosity in the maternal parent of the family B: unlike family A, family B showed absence of recombination among all the microsatellite markers that are within the *2La* region (data not shown).

While such uncertainties are inconsequential at present, they will need to be controlled if one attempts fine-

scale mapping. Eventually all the microsatellite markers should be mapped cytogenetically; this was not attempted here, because by the method used (hybridization with the entire clone insert) many clones hybridized to multiple sites, at least in part because of long simple sequence repeat arrays. Knowledge of the karyotypes of the actual parents would also be helpful, but this can only be done for the female at present (by examination of ovarian nurse cell polytene chromo-
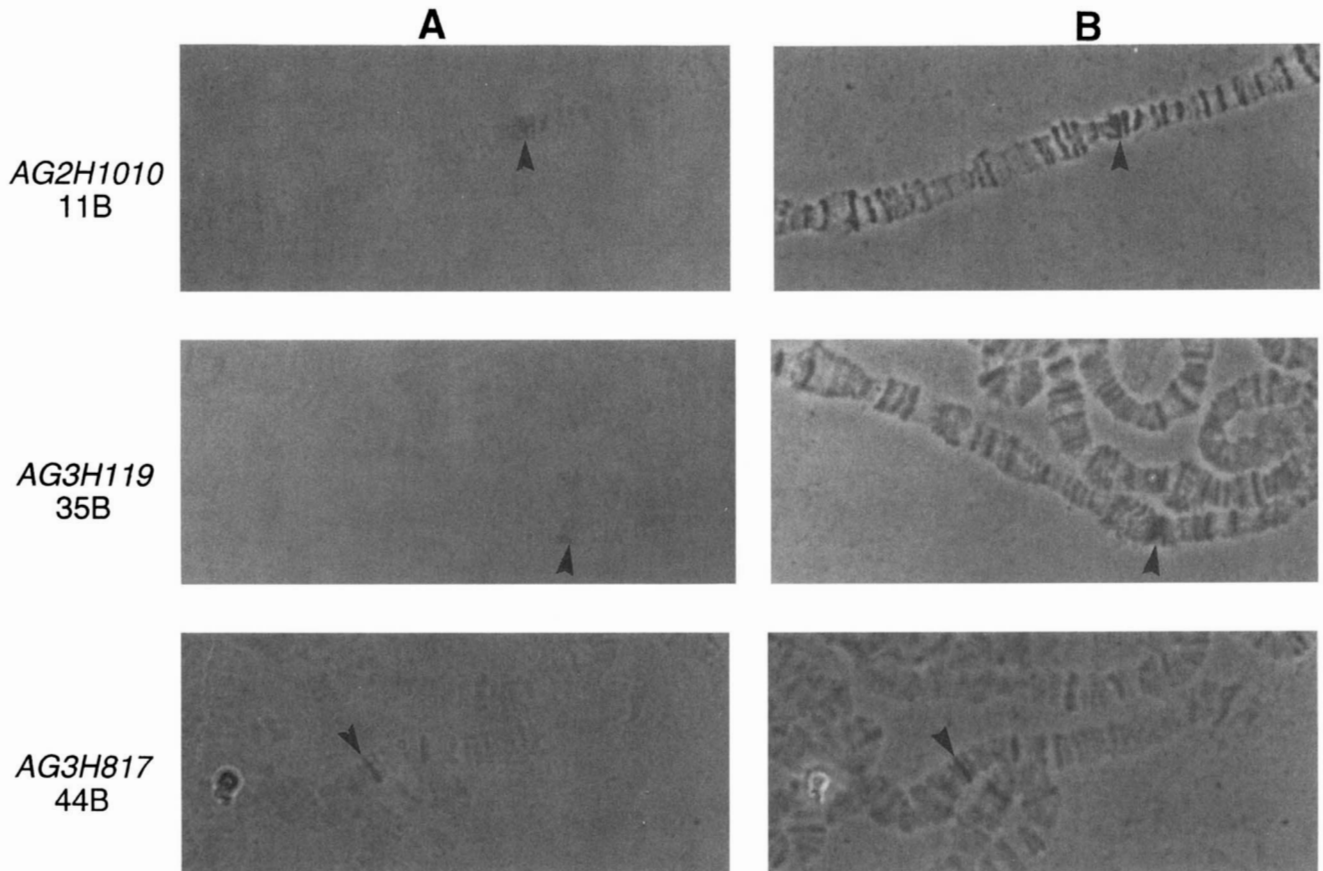
FIGURE 3.—Cytogenetic localization of microsatellite markers. The *AG2H1010* (top), *AG3H119* (middle), and *AG3H817* (bottom) markers were hybridized *in situ* to nurse cell polytene chromosomes. Bright-field images are on the left (A panels) and phase-contrast images on the right (B panels). Arrowhead indicates the site of hybridization.

somes); and if done, it limits the number of progeny that can be obtained. Molecular karyotyping by probes for inversion breakpoints would be valuable in this respect (DIMOPOULOS *et al.* 1996).

Despite these limitations, the present work represents a major advance in the genetics of *A. gambiae*. In an organism that initially had very few markers, this report makes available a large set of easily scorable markers that permit the mapping of genetic traits throughout most of the genome at a good average resolution. Because these markers are codominant, they allow the tracing of both homologous chromosomes in every generation. Because they are scored by PCR, these markers could all be scored in genotyping assays, each using less than 1/500th of the DNA from a single mosquito. Thus, additional microsatellite or other PCR-based markers can be easily introduced into the map, permitting finer scale mapping of areas of special interest. In this respect, it will be interesting to see how many microsatellites might be added to the third chromosome, which currently is more sparsely populated than the other two, by screening total and division-specific genomic DNA libraries more exhaustively, and by using oligonucleotides that represent several different simple sequence repeats.

Recently, microsatellite markers have been used extensively in genetic map construction (WEISSENBACH *et*

al. 1992), in genetic mapping of both quantitative and qualitative traits (ZHENG *et al.* 1993; LEVY-LAHAD *et al.* 1995), and in population (LANZARO *et al.* 1995) as well as evolutionary studies (BUCHANAN *et al.* 1994; GOLDSTEIN *et al.* 1995). Currently, the markers described here are being used to map the two different refractory mechanisms in *A. gambiae* (L. ZHENG, unpublished data). Some of these markers have also been tested and shown to be highly polymorphic in field-collected mosquitoes (LANZARO *et al.* 1995), potentially representing an extremely useful tool for field studies of population structure and dynamics. Finally, some of the markers can be scored in the sibling species, *A. arabiensis* (L. ZHENG, unpublished results), potentially facilitating studies of gene flow among these closely related and sympatric species.

REFERENCES

ASHBURNER, M., 1989 Mapping and exchange, pp 452–457 in *Drosophila, A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BEARD, C. B., M. Q. BENEDICT, J. P. PRIMUS, V. FINNERTY and F. H. COLLINS, 1995 Eye pigments in wild-type and eye-color mutant strains of the African malaria vector *Anopheles gambiae*. J. Hered. **86:** 375–380.

BENEDICT, M. Q., N. J. BESANSKY, H. CHANG, O. MUKABAYIRE and F. H. COLLINS, 1995 Mutations in the *Anopheles gambiae pink-eye* and *white* genes define distinct, tightly linked eye-color loci. J. Hered. **87:** 48–53.

BESANSKY, N. J., D. M. HAMM and F. H. COLLINS, 1994 An *Anopheles gambiae* cDNA predicts a protein similar to a yeast *Sui1* translation factor. Gene **141:** 299–300.

BESANSKY, N. J., J. A. BEDELL, O. MUKABAYIRE, D. HILFIKER and F. H. COLLINS, 1995 Cloning and characterization of the *white* gene from *Anopheles gambiae*. Insect Mol. Biol. **4:** 217–231.

BUCHANAN, F. C., L. J. ADAMS, R. P. LITTLEJOHN, J. F. MADDOX and A. M. CRAWFORD, 1994 Determination of evolutionary relationships among sheep breeds using microsatellites. Genomics **22:** 397–403.

CLEMENTS, A. N., 1992 Aspects of genetics, pp. 29–32 in *The Biology of Mosquitoes*, Vol. 1. Chapman and Hall, London.

COLLINS, F. H., R. K. SAKAI, K. D. VERNICK, S. PASKEWITZ, D. C. SEELEY *et al.*, 1986 Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. Science **234:** 607–610.

COLUZZI, M., and SABATINI, A., 1967 Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. Parassitologia **9:** 73–88

COLUZZI, M., A. SABATINI, V. PETRARCA and M. A. DI DECO, 1979 Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans. R. Soc. Trop. Med. Hyg. **73:** 483–497.

CREWS-OYEN, A. E., V. KUMAR and F. H. COLLINS, 1993 Association of two esterase genes, a chromosomal inversion, and susceptibility to *Plasmodium cynomolgi* in the African malaria vector *Anopheles gambiae*. Am. J. Trop. Med. Hyg. **49:** 341–347.

CURTIS, C. F., J. AKIYAMA and G. DAVIDSON, 1976 A genetic sexing system in *Anopheles gambiae* species A. Mosquito News **36:** 492–498.

DAVIDSON, G., 1956 Insecticide resistance in *Anopheles gambiae* Giles, a case of simple Mendelian inheritance. Nature **178:** 861–863.

DIMOPOULOS, G., L. ZHENG, V. KUMAR, A. DELLA TORRE, F. C. KAFATOS *et al.*, 1996 Integrated genetic map of *Anopheles gambiae*: use of RAPD polymorphisms for genetic, cytogenetic and STS landmarks. Genetics **143:** 953–960.

GOLDSTEIN, D. B., A. R. LINARES, L. L. CAVALLI-SFORZA and M. W. FELDMAN, 1995 Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. USA **92:** 6723–6727.

KILAMA, W. L., and G. B. CRAIG, JR., 1969 Monofactorial inheritance of susceptibility to *Plasmodium gallinaceum* in populations of *Aedes aegypti* (L.). J. Parasitol. **59:** 920–924.

KOSAMBI, D. D., 1944 The estimation of map distance from recombination values. Ann. Eugen. **12:** 172–175.

KNUDSON, D., L. ZHENG, S. W. GORDON, S. E. BROWN and F. C. KAFATOS, 1996 Genome organization of vectors, pp. 175–214 in *The Biology of Disease Vectors*, edited by B. J. BEATTY and W. C. MARQUARDT. Colorado State University Press, Fort Collins.

KUMAR, V. K., and F. H. COLLINS, 1994 A technique for nucleic acid *in situ* hybridization to polytene chromosomes of mosquitoes in the *Anopheles gambiae* complex. Insect Mol. Biol. **3:** 41–47.

LANDER, E., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. DALEY *et al.*, 1987 Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

LANZARO, G. C., L. ZHENG, Y. T. TOURE, S. F. TRAORE, F. C. KAFATOS *et al.*, 1995 Microsatellite DNA and isozyme variability in a West African population of *Anopheles gambiae*. Insect Mol. Biol. **4:** 105–112.

LEVY-LAHAD, E., E. M. WIJSMAN, E. NEMENS, L. ANDERSON, K. A. B. GODDARD *et al.*, 1995 A familial Alzheimer's disease locus on chromosome I. Science **269:** 970–973.

MACDONALD, W. W., 1962 The genetic basis of susceptibility to infection with semi-periodic *Brugia malayi* in *Aedes aegypti*. Ann. Trop. Med. Parasit. **56:** 373–382.

MARCHAND, R. P., 1984 Field observations on swarming and mating in *Anopheles gambiae* mosquitoes in Tanzania. Neth. J. Zool. **34:** 367–387.

MASON, G. F., and G. DAVIDSON, 1966 Morphological mutants in anopheline mosquitoes. Trans. R. Soc. Trop. Med. Hyg. **6:** 20.

MITCHELL, S. E., J. SEAWRIGHT and S. D. NARANG, 1993 Linkage map of the mosquito (*Anopheles quadrimaculatus* Species A), pp. 3.273–3.276 in *Genetic Maps*, Ed. 6, edited by S. J. O'BRIEN. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

MÜLLER, H.-M., J. M. CRAMPTON, A. DELLA TORRE, R. SINDEN and A. CRISANTI, 1993 Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. EMBO J. **12:** 2891–2900.

MÜLLER, H.-M., F. CATTERUCCIA, J. VIZIOLI, A. DELLA TORRE and A. CRISANTI, 1995 Constitutive and blood meal-induced trypsin genes in *Anopheles gambiae*. Exp. Parasitol. **81:** 371–385.

OLSON, M., L. HOOD, C. CANTOR and D. BOTSTEIN, 1989 A common language for physical mapping of the human genome. Science **245:** 1434–1435.

PASKEWITZ, S. M., and M. A. RIEHLE, 1994 Response of *Plasmodium*-refractory and susceptible strains of *Anopheles gambiae* to inoculated sephadex beads. Dev. Comp. Immunol. **18:** 369–375.

PATERSON, A., E. LANDER, S. LINCOLN, J. HEWITT, S. PETERSON *et al.*, 1988 Resolution of quantitative traits into Mendelian factors using a complete RFLP linkage map. Nature **335:** 721–726.

ROMANS, P., D. C. SEELEY, Y. KEW and R. W. GWADZ, 1991 Use of restriction fragment length polymorphism (RFLP) as a genetic marker in crosses of *Anopheles gambiae* (Diptera: Culicidae): independent assortment of a diphenol oxidase RFLP and an esterase locus. J. Med. Entomol. **28:** 147–151.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI *et al.*, 1988 Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:** 487–491.

SEAWRIGHT, J. A., and S. K. NARANG, 1993 Linkage map of the mosquito (*Anopheles albimanus*), pp. 3.269–3.272 in *Genetic Maps*, Ed. 6, edited by S. J. O'BRIEN. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SEVERSON, D. W., 1994 Applications of molecular marker analysis to mosquito vector competence. Parasitol. Today **10:** 336–340

SEVERSON, D. W., A. MORI, Y. ZHANG and B. M. CHRISTENSEN, 1993 Linkage map for *Aedes aegypti* using restriction fragment length polymorphisms. J. Hered. **84:** 241–247.

SEVERSON, D. W., A. MORI, Y. ZHANG and B. M. CHRISTENSEN, 1994 Chromosomal mapping of two loci affecting filarial worm susceptibility in *Aedes aegypti*. Insect Mol. Biol. **3:** 67–73.

SEVERSON, D. W., V. THATHY, A. MORI, Y. ZHANG and B. M. CHRISTENSEN, 1995 Restriction fragment length polymorphism mapping of quantitative trait loci for malaria susceptibility in the mosquito *Aedes aegypti*. Genetics **139:** 1711–1717.

STAM, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: Joinmap. Plant J. **5:** 739–744.

THATHY, V., D. W. SEVERSON and B. M. CHRISTENSEN, 1994 Reinterpretation of the genetics of susceptibility of *Aedes aegypti* to *Plasmodium gallinaceum*. J. Parasitol. **80:** 705–712.

VERNICK, K. D., and F. H. COLLINS, 1989 Association of a *Plasmodium*-refractory phenotype with an esterase locus in *Anopheles gambiae*. Am. J. Trop. Med. Hyg. **40:** 593–597.

VERNICK, K. D., F. H. COLLINS and R. W. GWADZ, 1989 A general system of resistance to malaria infection in *Anopheles gambiae* controlled by two main genetic loci. Am. J. Trop. Med. Hyg. **40:** 585–592.

VERNICK, K. D., H. FUJIOKA, D. C. SEELEY, B. TANDLER, M. AIKAWA *et al.*, 1995 *Plasmodium gallinaceum*: a refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae*. Exp. Parasitol. **80:** 583–595.

WEISSENBACH, J., G. GYAPAY, C. DIB, A. VIGNAL, J. MORESSETTE *et al.*, 1992 A second generation linkage map of the human genome. Nature **359:** 794–801.

ZHENG, L., R. D. C. SAUNDERS, D. FORTINI, A. DELLA TORRE, M. COLUZZI *et al.*, 1991 Low-resolution genome map of the malaria mosquito *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA **88:** 11187–11191.

ZHENG, L., F. H. COLLINS, V. KUMAR and F. C. KAFATOS, 1993 A detailed genetic map for the X chromosome of the malaria vector, *Anopheles gambiae*. Science **261:** 605–608.

Communicating editor: V. G. FINNERTY